Stochastic Processes and their Applications 14 (1983) 157-174
North-Holland Publishing Company

APPROXIMATIONS OF SOME HAZARD RATE ESTIMATORS IN A COMPETING RISKS MODEL

Murray D. BURKE*

Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta T2N 1N4, Canada

Received 14 August 1980
Revised 2 October 1981

A general model involving k competing risks is studied and the hazard rates of these risks are simultaneously estimated. The estimators are strongly approximated by Gaussian processes and the limiting distribution of certain statistics are obtained.

Hazard rates	competing risks
strong approximations	Gaussian processes
Density estimation	

1. Introduction

There has been much interest in the log survivor function and the hazard rate function. They are utilized in reliability studies [7], studies of mortality [14] and in seismology [12, 26]. Watson and Leadbetter [27, 28] and Rice and Rosenblatt [23] have studied a number of estimators for these functions based on a single random sample. In the present exposition we shall study a more general competing risks model in which the hazard rate for k risks are simultaneously estimated. Strong approximations of these estimates in terms of Gaussian processes are obtained.

Our model can be defined as follows: Let X be a real-valued random variable with a continuous distribution function $F(x)=\mathbf{P}\{X<x\}$. For a fixed natural number k, let $A^1, A^2, \ldots, A^k$ be pairwise disjoint sets, (or at least $\mathbf{P}(A^i \cap A^m)=0$ for $i \neq m$), such that $\mathbf{P}(\bigcup_{i=1}^k A^i)=1$ and define the sub-distribution function $\tilde{F}_i(x)=\mathbf{P}\{X<x$ and $A^i\}$, $i=1,2,\ldots,k$. When observing X we are interested in the joint behavior of the pairs (X, A^i). In survival analysis terminology X corresponds to the survival time of an individual who dies from one of k causes $A^1, A^2, \ldots, A^k$. In reliability terminology X may correspond to the lifetime of a system which fails from one of k causes $A^1, A^2, \ldots, A^k$. We are assuming that A^i and A^m cannot occur simultaneously, $i \neq m$.

* Research partially supported by a Canadian NSERC grant.

0304-4149/83/0000-0000/$03.00

The following definition of hazard (or mortality) rate in the present model is useful: *Given that an individual (or system) has survived all risks $A^1, A^2, \ldots, A^k$ up to time x the instantaneous probability rate that he/she dies from specific cause A^i at time x is the hazard rate for risk A^i at time x.* On assuming that the sub-density functions $\tilde{f}_i(x) = \tilde{F}_i'(x)$ exist, the *hazard rate* for risk A^i at time t_i is defined by

$$h_i(t_i) = \tilde{f}_i(t_i)/(1 - F(t_i)). \tag{1.1}$$

We define the hazard rate vector H by

$$H(t) = (h_1(t_1), h_2(t_2), \ldots, h_k(t_k)) \tag{1.2}$$

where $t = (t_1, t_2, \ldots, t_k) \in \mathbb{R}^k$. Let $J \subset \{1, 2, \ldots, k\}$ be non-empty. Then the hazard rate for the combined risks $\bigcup_{i \in J} A^i$ is obtained as $\sum_{i \in J} h_i(x)$. The total hazard rate for an individual or system is then $\sum_{i=1}^k h_i(x)$.

Based on a sequence of independent replicas $\{X_j, A_j^1, A_j^2, \ldots, A_j^k\}$, $j = 1, 2, \ldots$, each having the same distribution as $\{X, A^1, A^2, \ldots, A^k\}$, we wish to estimate the hazard rate vector H of (1.2). We will employ the popular kernel method. There are a number of excellent survey papers on curve estimates (cf. for example, [24]).

Let F_n denote the empirical distribution function based on $X_1, X_2, \ldots, X_n$,

$$F_n(x) = n^{-1} \sum_{j=1}^{n} I\{X_j < x\}, \quad x \in \mathbb{R}$$

where $I\{A\}$ is the indicator function of the event A. Let w_i, our kernel, be any density function satisfying conditions (A2) of Section 3. We estimate the sub-density function $\tilde{f}_i$ by the empirical sub-density,

$$\tilde{f}_{in}(t_i) = (nb_n)^{-1} \sum_{j=1}^{n} \delta_j^i w_i((t_i - X_j)/b_n) = \int b_n^{-1} w_i((t_i - u)/b_n)\, \mathrm{d}\tilde{F}_{in}(u), \tag{1.3}$$

where $\tilde{F}_{in}$ is the empirical sub-distribution function,

$$\tilde{F}_{in}(u) = n^{-1} \sum_{j=1}^{n} I\{X_j \le u \text{ and } A_j^i\}, \tag{1.4}$$

and δ_j^i is the indicator function of A_j^i. Thus, we obtain the estimator

$$h_{in}^{(1)}(t_i) = \tilde{f}_{in}(t_i)/(1 - F_n(t_i)) \tag{1.5}$$

for $h_i(t_i)$ and

$$H_n^{(1)}(t) = (h_{1n}^{(1)}(t_1), h_{2n}^{(1)}(t_2), \ldots, h_{kn}^{(1)}(t_k)) \tag{1.6}$$

for $H(t)$, where $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$.

An alternate estimator for h_i is

$$\begin{aligned} h_{in}^{(2)}(t_i) &= (nb_n)^{-1} \sum_{j=1}^{n} \delta_j^i w_i((t_i - X_j)/b_n)(1 - F_n(X_j)) \\ &= \int b_n^{-1} w_i((t_i - u)/b_n)(1 - F_n(u))^{-1}\, \mathrm{d}\tilde{F}_{in}(u) \end{aligned} \tag{1.7}$$

and hence we can estimate $H(t)$ by

$$H_n^{(2)}(t) = (h_{1n}^{(2)}(t_1), h_{2n}^{(2)}(t_2), \ldots, h_{kn}^{(2)}(t_k)).$$

Although, in Section 5, we show that the estimators (1.5) and (1.7) are asymptotically equivalent, one may be preferred by the statistician for small and moderate sample sizes. However, a detailed comparison of these estimators is beyond the scope of the present exposition.

In Section 3 a strong approximation theorem is proved for the sub-densities $(\tilde{f}_{1n}, \tilde{f}_{2n}, \ldots, \tilde{f}_{kn})$ and is used in the approximation of $H_n^{(1)}$ by Gaussian processes in Section 4. The asymptotic distribution of a global measure of deviation as well as a strong law are obtained. These results are useful for constructing confidence bands and goodness-of-fit tests. Section 5 is concerned with showing that $H_n^{(1)}$ and $H_n^{(2)}$ are asymptotically equivalent under certain regularity conditions. Some further results are discussed in Section 6 including the case when parameters of the underlying hazard rates are estimated. We begin by discussing an alternate approach to our model.

2. The model: an alternate approach

Despite some criticism in the literature, for example in [20], the following approach to the model is often illuminating. For a fixed natural number k, let $(Y_j^1, Y_j^2, \ldots, Y_j^k)$, $j = 1, 2, \ldots$, be a sequence of independent and identically distributed random vectors having continuous joint survival distribution given by

$$S(y_1, y_2, \ldots, y_k) = \mathbf{P}\{Y_j^1 > y_1, Y_j^2 > y_2, \ldots, Y_j^k > y_k\} \tag{2.1}$$

where $(y_1, y_2, \ldots, y_k) \in \mathbb{R}^k$. What one observes is the sequence $(X_j, \delta_j^1, \delta_j^2, \ldots, \delta_j^k)$, $j = 1, 2, \ldots$, where

$$X_j = \min\{Y_j^1, Y_j^2, \ldots, Y_j^k\}$$

and $\delta_j^i = I\{X_j = Y_j^i\}$ is the indicator function of the event $A_j^i = \{X_j = Y_j^i\}$ for $i = 1, 2, \ldots, k$. Since S is assumed to be continuous in this paper, all but one δ_j^i $(i = 1, 2, \ldots, k)$ are equal to zero.

The X_j have distribution function given by

$$F(x) = 1 - S(x, x, \ldots, x), \quad x \in \mathbb{R}.$$

We will assume that the gradient function $\nabla S(y_1, y_2, \ldots, y_k)$ having ith component $(\partial/\partial y_i)S(y_1, y_2, \ldots, y_k)$ exists at all $(y_1, y_2, \ldots, y_k) = (x, x, \ldots, x)$, $x \in \mathbb{R}$ and let

$$\tilde{f}_i(x) = -(\partial/\partial y_i)S(y_1, y_2, \ldots, y_k)|_{\text{all } y_i = x},$$

where the partial derivative is evaluated at $y_1 = y_2 = \cdots = y_k = x$.

We have

$$\tilde{F}_i(x) = \mathbf{P}\{X_j \leq x, \delta_j^i = 1\} = \int_{-\infty}^{x} \tilde{f}_i(v)\, dv.$$

On letting $h_i(x) = \tilde{f}_i(x)/(1-F(x))$, we arrive at the model of Section 1. Note that we have not assumed any independence condition on the components of $(Y_j^1, Y_j^2, \ldots, Y_j^k)$. The results obtained in subsequent sections are valid without such a condition.

Examples. (1) In reliability theory one may consider a system j consisting of k component sub-systems connected in series. The whole system fails when one of the sub-systems fails. On letting Y_j^i denote the lifetime of the ith component of system j, the failure time of the whole system is $X_j = \min\{Y_j^1, \ldots, Y_j^k\}$ and δ_j^i ($i = 1, 2, \ldots, k$) indicates if the system failure is due to the failure of the ith component.

(2) In survival analysis Y_j^i corresponds to the survival time of individual j until he/she dies from cause i. Each individual is subject to k competing risks and the (theoretical) survival time for risk i may be right censored by another risk from which the individual dies. Again, one only observes X_j and δ_j^i ($i = 1, 2, \ldots, k$).

Remark. Langberg, Proschan and Quinzi [17] have shown that, under our conditions on X_j and δ_j^i in Section 1, one can construct an independent sequence $(L_j^1, L_j^2, \ldots, L_j^k)$ of random vectors such that if

$$X_j^* = \min\{L_j^1, L_j^2, \ldots, L_j^k\} \quad \text{and} \quad \delta_j^{i*} = I\{X_j^* = L_j^i\},$$

then

$$\{X_j^*, \delta_j^{i*}: i = 1, 2, \ldots, k\} \overset{D}{=} \{X_j, \delta_j^i; i = 1, 2, \ldots, k\},$$

that is, their joint distributions are equal. In fact, their construction yields independent components $L_j^1, L_j^2, \ldots, L_j^k$. Thus, there is a direct correspondence between the approaches of Sections 1 and 2. The L_j^i may not be the same as the original Y_j^i. This latter result also shows that with the observations $(X_j, \delta_j^1, \ldots, \delta_j^k)$ having distribution functions F and $\tilde{F}_i$, one is not able to distinguish between the independent risks model with 'latent' lifetimes $(L_j^1, \ldots, L_j^k)$ and an infinite number of dependent risks models with dependent latent lifetimes $(Y_j^1, \ldots, Y_j^k)$. This has been observed previously by Cox [6], Tsiatis [25] and Peterson [19]. In mortality studies it is unrealistic to expect that all the risks from various diseases are independent.

Let $J \subset \{1, 2, \ldots, k\}$ be non-empty. One is sometimes interested in making inference about the sub-system $\{Y_j^i\}_{i \in J}$ where J^c is the relative complement of J with respect to $\{1, 2, \ldots, k\}$.

Let

$$F_J(x) = \mathbf{P}\{Y_j^i < x, \text{for all } i \in J\} \tag{2.2}$$

and $f_J = F_J'$. The *marginal hazard rate* for the sub-system J is

$$h_J(x) = f_J(x)/(1 - F_J(x)). \tag{2.3}$$

Langberg, Proschan and Quinzi [17] have obtained conditions whereby F_i can be consistently estimated based on the observations $\{X_j, \delta_j^1, \ldots, \delta_j^k\}$, $j = 1, 2, \ldots$. As pointed out by Arjas [2], assuming our continuity assumptions, their conditions are equivalent to the equality

$$\sum_{i \in J} h_i(x) = h_J(x), \tag{2.4}$$

where h_i is defined by (1.1). Thus, to estimate the hazard rate h_J, assuming (2.4), one can use the estimators (1.5) or (1.7).

However, it should be noted that condition (2.4) cannot be tested based on the data $\{X_j, \delta_j^1, \ldots, \delta_j^k\}$, $j = 1, 2, \ldots$. It is an a priori assumption based on the investigator's knowledge of the system. Essentially (2.4) is an independence statement of sub-system J from sub-system J^c at time x. Prentice et al. [20] argue convincingly that in survival analysis (2.4) is not generally satisfied since diseases usually have interactive effects. Still, risks such as accidental death can often be assumed to be independent from disease-related risks and can be viewed as a random censoring of the disease-related risks of interest. In this case the present approach of assuming (2.4) and using (1.5) and (1.7) to estimate the marginal hazard rate h_J, where J^c consists of the independent censoring risks, is pertinent. In general, the reliability analyst is in more of a position to make a priori assumptions like (2.4) than the bio-medical investigator.

3. Asymptotic behaviour of $\tilde{f}_{in}$

In order to obtain our approximation results we will assume the following conditions for each i, ($i = 1, 2, \ldots, k$):

(A1) (i) The sub-density function $\tilde{f}_i$ is positive in the interval (α_i, β_i), $-\infty \leq \alpha_i < \beta_i \leq \infty$ and vanishes outside.

(ii) $\tilde{f}_i$ has a bounded second derivative $\tilde{f}_i''$ (and hence $\tilde{f}_i$ and $\tilde{f}'$ are also bounded).

(A2) (i) w_i is a bounded density function and either (a) is zero outside an interval $[-A_i, A_i]$ and absolutely continuous on $[-A_i, A_i]$ with finite limits $w_i(A_i) = \lim_{x \uparrow A_i} w_i(x)$ and $w_i(-A_i) = \lim_{x \downarrow -A_i} w_i(x)$, or (b) is absolutely continuous on $(-\infty, \infty)$ with $w_i' \in L^1$ and L^2.

(ii) $\int_{-\infty}^{\infty} x w_i(x)\,dx = 0$, $\int_{-\infty}^{\infty} x^2 w_i(x)\,dx < \infty$.

(A3) (i) $\{b_n\}$ is a sequence of positive reals such that $b_n \downarrow 0$ and $nb_n \to \infty$.

(ii) $(\alpha_{in}, \beta_{in})$ is a sequence of intervals such that $(\alpha_{in}, \beta_{in}) \subset (\alpha_{i,n+1}, \beta_{i,n+1})$, $(\alpha_{in}, \beta_{in}) \to (\alpha_i, \beta_i)$ and $\log(\beta_{in} - \alpha_{in}) = O(n$. Let $a_n = \max_{1 \leq i \leq k} \sup_{\alpha_{in} \leq x \leq \beta_{in}} (\tilde{f}_i(x))^{-1/2}$ and assume

$$b_n[1 - F(\beta_{in})]^{-1} \to 0 \quad \text{and} \quad a_n b_n^2 \to 0 \qquad \text{as } n \to \infty.$$

Remarks. (1) Assumptions (A1), (A2) and (A3) are fairly standard. (A1) concerns the existence of second derivatives which are needed for Lemma 3.1. (A2) concerns the choice of a 'nice' symmetric density w_i for our kernel, while (A3) discusses the choice of the bandwidth sequence b_n and also how the interval $(\alpha_{in}, \beta_{in})$ over which our approximations are made fills out the whole support of $\tilde{f}_i$. The kernel estimates depend a geat deal on both the kernel and the bandwidth chosen. An 'optimal' weight function as been obtained by Epanechnikov [10], namely $w_i(u) = 3/(4(5)^{1/2})(1-(u^2/5))$ if $|u| \leq 5^{1/2}$ and equals zero otherwise.

(2) Note that if $\tilde{f}_i$ is bounded away from zero on the finite interval (α_i, β_i), then $a_n = O(1)$. In this case, the rates of convergence in our theorems below will be much faster than what is stated in the remarks following Theorems 3.2, 4.1 and 5.1. For example, the δ in $\{b_n\}$ can satisfy $\frac{1}{5} < \delta < \frac{1}{2}$ and convergence can take place over (α_i, β_{in}) for Theorems 4.1 and 5.1 and over (α_i, β_i) for Theorem 3.2.

Let us define the empirical processes

$$\tilde{\alpha}_{in}(x) = n^{1/2}[\tilde{F}_{in}(x) - \tilde{F}_i(x)], \quad i = 1, 2, \ldots, k$$

and

$$\alpha_{0n}(x) = n^{1/2}[F_n(x) - F(x)]$$

where $x \in \mathbb{R}$ and $\tilde{F}_{in}$, $\tilde{F}_i$, F_n and F are defined as in Section 1. The following result is crucial.

Theorem 3.1. (Burke, Csörgö and Horváth [5]. *On a suitable probability space, one can define* $k+1$ *two-parameter Gaussian processes* $K_0(x, n)$, $\tilde{K}_1(x, n), \ldots, \tilde{K}_k(x, n)$ *such that we have for* $\alpha_n(t) = (\alpha_{0n}(t_0), \tilde{\alpha}_{1n}(t_1), \ldots, \tilde{\alpha}_{kn}(t_k))$ *and* $V(t, n) = (K_0(t_0, n), \tilde{K}_1(t_1, n), \ldots, \tilde{K}_k(t_k, n))$, $t = (t_0, t_1, \ldots, t_k)$,

$$\sup_{t \in R^{k+1}} \|\alpha_n(t) - n^{-1/2} V(t, n)\| \stackrel{\text{a.s.}}{=} O(n^{-1/2} \log^2 n),$$

where $V(t, n)$ *itself is a* $(k+1)$*-dimensional vector-valued Gaussian process having the same covariance structure as the vector* $n^{1/2}\alpha_n(t)$, *namely* $\mathbf{E}\, V(t, n) = 0$, *the zero vector, and for any* $i, j = 1, 2, \ldots, k$, $i \neq j$,

$$\begin{aligned}
&\mathbf{E}\, K_0(x, n) K_0(y, m) = (n \wedge m)\{F(x) \wedge F(y) - F(x)F(y)\}, \\
&\mathbf{E}\, \tilde{K}_i(x, n) \tilde{K}_i(y, m) = (n \wedge m)\{\tilde{F}_i(x) \wedge \tilde{F}_i(y) - \tilde{F}_i(x)\tilde{F}_i(y)\}, \\
&\mathbf{E}\, \tilde{K}_i(x, n) \tilde{K}_j(y, m) = (n \wedge m)\{-\tilde{F}_i(x)\tilde{F}_j(y)\}, \\
&\mathbf{E}\, \tilde{K}_i(x, n) K_0(y, m) = (n \wedge m)\{\tilde{F}_i(x) \wedge \tilde{F}_i(y) - \tilde{F}_i(x)F(y)\}.
\end{aligned} \tag{3.1}$$

where $a \wedge b = \min\{a, b\}$.

Remarks. (1) The phrase "a suitable probability space" is meant in the sense of Komlós, Major and Tusnády [15], on which Theorem 3.1 is based. The processes

$K_0, \tilde{K}_1, \ldots, \tilde{K}_k$ are Kiefer processes. A Kiefer process is a separable Gaussian process defined on $[0,1]\times[0,\infty)$ such that $\mathbf{E}\,K(s,y)=0$ and

$$\mathbf{E}\,K(s_1,y_1)K(s_2,y_2)=(y_1\wedge y_2)\{s_1\wedge s_2-s_1s_2\}. \tag{3.2}$$

(2) Theorem 3.1 can also be stated in terms of sequences of Brownian bridges $B_{0n}, \tilde{B}_{1n}, \ldots, \tilde{B}_{kn}$ approximating α_n at the rate of convergence $O(n^{-1/2}\log n)$. By using the Kiefer process representation one can deduce strong laws for $H_n^{(1)}$ and $H_n^{(2)}$ (cf. Theorem 4.2), as well as weak convergence of statistics based on $H_n^{(1)}$ and $H_n^{(2)}$.

Consider the process

$$T_{in}(x)=(nb_n/\tilde{f}_i(x))^{1/2}(\tilde{f}_{in}(x)-\tilde{f}_i(x)),\quad x\in(\alpha_i,\beta_i). \tag{3.3}$$

Using Theorem 3.1, we wish to strongly approximate T_{in}. If α_i or β_i are finite, we will assume that ω_i has finite support and obtain the following lemma.

Lemma 3.2. *Under conditions* (A1), (A2) *and* (A3)(i),

$$\sup_{\alpha_i<x<\beta_i}|\mathbf{E}\,\tilde{f}_{in}(x)-\tilde{f}_i(x)|\leq Cb_n^2.$$

For a proof, we refer to [21, Lemma 6a]. Note that the expected value

$$\mathbf{E}\,\tilde{f}_{in}(x)=\int b_n^{-1}w_i(x-u)/b_n)\,\mathrm{d}\tilde{F}_i(x).$$

Consequently, assuming also condition (A3)(ii) to hold,

$$\sup_{\alpha_{in}<x<\beta_{in}}(nb_n/\tilde{f}_i(x))^{1/2}|\mathbf{E}\,\tilde{f}_{in}(x)-\tilde{f}_i(x)|=O(r_1(n)) \tag{3.4}$$

where $r_1(n)=n^{1/2}b_n^{5/2}a_n$ and a_n is defined in (A3)(iii).

If $r_1(n)=o(1)$, then T_{in} is asymptotically equivalent to

$$\begin{aligned}T^*_{in}(x)&=(nb_n/\tilde{f}_i(x))^{1/2}(\tilde{f}_{in}(x)-\mathbf{E}\,\tilde{f}_{in}(x))\\&=(b_n\tilde{f}_i(x))^{-1/2}\int w_i(x-u)/b_n)\,\mathrm{d}\tilde{\alpha}_{in}(u)\end{aligned} \tag{3.5}$$

where α_{in} is defined as in Theorem 3.1. As in [3], using Theorem 3.1, we obtain

$$\sup_{\alpha_{in}<x<\beta_{in}}|T^*_{in}(x)-G^*_i(x,n)|\overset{\text{a.s.}}{=}O(r_2(n)) \tag{3.6}$$

where $r_2(n)=(nb_n)^{-1/2}(\log n)^2a_n$ and G^*_i is the Gaussian process

$$G^*_i(x,n)=(b_n\tilde{f}_i(x))^{-1/2}\int w_i((x-u)/b_n)\,\mathrm{d}n^{-1/2}\tilde{K}_i(u,n).$$

Further, we note that the Kiefer process $\tilde{K}_i(u,n)$ has representation

$$\tilde{K}_i(u,n)=\tilde{W}_i(\tilde{F}_i(u),n)-\tilde{F}_i(u)\tilde{W}_i(1,n) \tag{3.7}$$

where $\tilde{W}_i$ is a two-parameter Wiener process, that is, a Gaussian process with $\mathbf{E}\,\tilde{W}_i(s,n)=0$ and $\mathbf{E}\,\tilde{W}_i(s,n)\tilde{W}_i(u,m)=(n\wedge m)(s\wedge u)$. We have

$$\sup_{\alpha_{in}<x<\beta_{in}} (b_n\tilde{f}_i(x))^{-1/2}\left|\int w_i((x-u)/b_n)\,\mathrm{d}\tilde{F}_i(u)n^{-1/2}W_i(1,n)\right| \overset{\text{a.s.}}{=} \mathrm{O}(r_3(n)) \tag{3.8}$$

where $r_3(n)=(b_n\log\log n)^{1/2}$. This follows from

$$[\tilde{f}_i(x)]^{-1/2}\,\mathbf{E}\,\tilde{f}_{in}(x)\leqslant(\tilde{f}_i(x))^{1/2}+[\tilde{f}_i(x)]^{-1/2}Cb_n^2 \quad\text{and}\quad a_nb_n^2\to 0.$$

We obtain the following.

Theorem 3.3. *Assuming conditions* (A1), (A2) *and* (A3) *one can construct k independent two-parameter Gaussian processes G_i, $1\leqslant i\leqslant k$, such that for $t=(t_1,t_2,\ldots,t_k)$*

$$\sup_{t\in R(n)}\|T_n(t)-G(t,n)\| \overset{\text{a.s.}}{=} \mathrm{O}\Big(\sup_{1\leqslant i\leqslant 3} r_i(n)\Big) \tag{3.9}$$

where $R(n)$ is the rectangle $\prod_{i=1}^k(\alpha_{in},\beta_{in})$,

$$T_n(t)=(T_{1n}(t_1),T_{2n}(t_2),\ldots,T_{kn}(t_k))$$

and

$$G(t,n)=(G_1(t_1,n),G_2(t_2,n),\ldots,G_k(t_k,n)),$$

with T_{in} defined by (3.5) *and*

$$G_i(t_i,n)=(b_n\tilde{f}_i(t_i))^{-1/2}\int\omega_i((t_i-u)/b_n)\,\mathrm{d}n^{-1/2}W_i(\tilde{F}_i(u),n).$$

Remark. If $b_n=n^{-\delta}$ with $\frac{2}{5}<\delta<\frac{1}{2}$ and $a_n=n^{(1-\delta-\varepsilon)/2}$ for some $\varepsilon>0$ with $\delta-\varepsilon>0$, then $\max_{1\leqslant i\leqslant n} r_i(n)=\mathrm{O}(n^{-\varepsilon'})$ for some $\varepsilon'>0$.

Proof of Theorem 3.3. The approximation (3.9) follows from (3.4), (3.6) and (3.8). What is also of interest is the assertion that the components G_i of the vector-valued Gaussian process G are independent. To prove this we must refer to the proof of Theorem 3.1 (cf. [5, Theorem 3.1]), where the Kiefer processes $\tilde{K}_i$ are constructed. The Kiefer processes $\tilde{K}_i$ are not independent, having nonzero covariance structure (3.1). However, we will show that the Wiener processes $\tilde{W}_i$, defined in (3.7), are independent.

The $\tilde{K}_i$ are defined in terms of a single Kiefer process K having covariance structure (3.2). Let $\tilde{F}_i(\infty)=\lim_{x\to\infty}\tilde{F}_i(x)$. Then, $\tilde{F}_1(\infty)+\tilde{F}_2(\infty)+\cdots+\tilde{F}_k(\infty)=1$

and for $u \in \mathbb{R}$

$$\begin{aligned}
\tilde{K}_1(u, n) &= K(\tilde{F}_1(u), n), \\
\tilde{K}_2(u, n) &= K(\tilde{F}_2(u) + \tilde{F}_1(\infty), n) - K(\tilde{F}_1(\infty), n), \\
&\vdots \\
\tilde{K}_i(u, n) &= K(\tilde{F}_i(u) + \tilde{F}_1(\infty) + \tilde{F}_2(\infty) + \cdots + \tilde{F}_{i-1}(\infty), n) \\
&\qquad - K(\tilde{F}_1(\infty) + \cdots + \tilde{F}_{i-1}(\infty), n)
\end{aligned}$$

for $i = 1, 2, \ldots, k$. The Kiefer process K has representation in terms of a two-parameter Wiener process W,

$$K(y, n) = W(y, n) - yW(1, n), \quad 0 \leq y \leq 1, n \geq 0.$$

Hence, the Wiener process $\tilde{W}_i$ of (3.7) has representation

$$\begin{aligned}
\tilde{W}_1(\tilde{F}_1(u), n) &= W(\tilde{F}_1(u), n), \\
\tilde{W}_2(\tilde{F}_2(u), n) &= W(\tilde{F}_2(u) + \tilde{F}_1(\infty), n) - W(\tilde{F}_1(\infty), n), \\
&\vdots \\
\tilde{W}_i(\tilde{F}_i(u), n) &= W(\tilde{F}_i(u) + \tilde{F}_1(\infty) + \tilde{F}_2(\infty) + \cdots + \tilde{F}_{i-1}(\infty), n) \\
&\qquad - W(\tilde{F}_1(\infty) + \cdots + \tilde{F}_{i-1}(\infty), n)
\end{aligned}$$

for $i = 1, 2, \ldots, k$. Thus, $(W_1, W_2, \ldots, W_k)$ is a vector-valued Gaussian process (being defined in terms of the same W), consisting of Wiener processes and it is easy to verify that these Wiener processes are independent.

As in [23, Theorem 5] we obtain the following.

Corollary 3.4. *Assume conditions* (A1), (A2) *and* (A3) *to hold. Let* $b_n = n^{-\delta}$, *where* $\frac{1}{5} < \delta < \frac{1}{2}$ *and choose the sequence* $(\alpha_{in}, \beta_{in})(1 \leq i \leq k)$ *so that*

$$a_n = \max_{1 \leq i \leq k} \sup_{\alpha_{in} < t_i < \beta_{in}} (f_i(t_i))^{-1/2} = \mathrm{O}(n^{(1-\delta-\varepsilon)/2})$$

and

$$\sup_{\alpha_{in} < t_i < \beta_{in}} |\tilde{f}_i'(t_i)/\tilde{f}_i(t_i)| = \mathrm{O}(n^{(\delta-\varepsilon)/2})$$

as $n \to \infty$, *for some* $\varepsilon > 0$ *with* $1 - \delta - \varepsilon, \delta - \varepsilon > 0$. *Let*

$$c_{in} = (\beta_{in} - \alpha_{in}) b_n^{-1}.$$

Then

$$\lim_{n \to \infty} \mathbf{P}\{M_n - d_n < (y_1, y_2, \ldots, y_k)\} = \exp\left\{ -2 \sum_{i=1}^{k} \mathrm{e}^{-y_i} \right\}$$

where

$$M_n = \Big((2\log c_{1n})^{1/2} \sup_{\alpha_{1n}<t_1<\beta_{1n}} |T_{1n}(t_1)|(\lambda(w_1))^{-1/2}, \ldots, (2\log c_{kn})^{1/2}$$

$$\times \sup_{\alpha_{kn}<t_k<\beta_{kn}} |T_{kn}(t_k)|(\lambda(w_k))^{-1/2}\Big),$$

T_{in} is defined by (3.3), $\lambda(w_i) = \int [w_i(u)]^2\, du$ *and* $d_n = (d_{1n}, d_{2n}, \ldots, d_{kn})$ *with*

$$d_{in} = (2\log c_{in})^{1/2} + (2\log c_{in})^{-1/2}\{\log \pi^{-1/2} K_1(w_i) + \tfrac{1}{2}\log\log c_{in}\},$$

and $K_1(w_i) = (w_1^2(A_i) + w_i^2(-A_i))/(2\lambda(w_i))$ *if* $K_1(w_i) > 0$, *and otherwise*

$$d_{in} = (2\log c_{in})^{1/2} + (2\log c_{in})^{-1/2}(2\pi)^{-1} K_2(w_i)\}$$

where $K_2(w_i) = \{\int [w_i'(u)]^2\, du/\lambda(w_i)\}^{1/2}$.

4. Asymptotic behaviour of $H_n^{(1)}$

Consider the processes $\bar{T}_n(t) = (\bar{T}_{1n}(t_1), \bar{T}_{2n}(t_2), \ldots, \bar{T}_{kn}(t_k))$ defined by

$$\bar{T}_{in}(t_i) = (nb_n/\bar{f}_i(t_i))^{1/2}(1-F(t_i))(h_{in}^{(1)}(t_i) - h_i(t_i)). \tag{4.1}$$

On noting that

$$\begin{aligned} h_{in}^{(1)}(x) - h_i(x) &= (1-F(x))^{-1}(\bar{f}_{in}(x) - \bar{f}_i(x)) \\ &\quad - \bar{f}_{in}(x)\{(1-F(x))^{-1} - (1-F_n(x))^{-1}\} \\ &= (1-F(x))^{-1}(\bar{f}_{in}(x) - \bar{f}_i(x)) \\ &\quad - \bar{f}_{in}(x)\{n^{-1/2}\alpha_{0n}(x)(1-F_n(x))^{-1}(1-F(x))^{-1}\} \end{aligned}$$

and since

$$\sup_{x<\beta_{in}} (1-F_n(x))^{-1} \overset{\text{a.s.}}{=} O(1-F(\beta_{in}))^{-1} \quad \text{for } 1-F(\beta_{in}) \geq \left(2(1+\delta)\frac{\log n}{n}\right)$$

(cf. [5, Lemma 4.1]) we obtain

$$\sup_{\alpha_{in}<x<\beta_{in}} |(1-F(x))(h_{in}^{(1)}(x) - h(x)) - (\bar{f}_{in}(x) - \bar{f}_i(x))| \overset{\text{a.s.}}{=} O(r_4(n))$$

where $r_4(n) = n^{-1/2}(\log\log n)^{1/2}(1-F(\beta_{in}))^{-1}$. Hence we have the following theorem.

Theorem 4.1. *Under the conditions of Theorem 3.3, one can construct k independent two-parameter Gaussian processes G_i such that, for $t = (t_1, \ldots, t_k)$,*

$$\sup_{t\in R(n)} \|\bar{T}_n(t) - G(t, n)\| \overset{\text{a.s.}}{=} O\Big(\max_{1\le i\le 4} r_i(n)\Big) \tag{4.2}$$

where $\bar{T}_n(t)$ *is defined by* (4.1), $R(n)=\prod_{i=1}^{k}(\alpha_{in},\beta_{in})$,

$$1-F(\beta_{in})\geq[2(1+\delta)n^{-1}\log n]^{1/2}$$

and $G(t,n)=(G_1(t_1,n),\ldots,G_k(t_k,n))$ *is defined as in Theorem* 3.3.

Moreover, if the conditions of Corollary 3.4 *hold, we obtain*

$$\lim_{n\to\infty}\mathbf{P}\{\bar{M}_n-d_n<(y_1,y_2,\ldots,y_k)\}=\exp\left\{-2\sum_{i=1}^{k}\mathrm{e}^{-y_i}\right\}$$

where $\bar{M}_n=(\bar{M}_{1n},\bar{M}_{2n},\ldots,\bar{M}_{kn})$ *with*

$$\bar{M}_{in}=(2\log c_{in})^{1/2}\sup_{\alpha_{in}<t<\beta_{in}}|\bar{T}_{in}(t_i)|,$$

and c_{in} *and* d_n *are defined as in Corollary* 3.4.

If b_n and a_n are as in Corollary 3.4 and if $(1-F(\beta_{in}))^{-1}=\mathrm{O}(n^{1/2-\lambda})$ for some $\lambda>0$, then $\max_{1\leq i\leq 4}r_i(n)=\mathrm{O}(n^{-\varepsilon'})$ for some $\varepsilon'>0$.

If we assume that each $\tilde{f}_i$ is bounded away for zero on the finite interval (α_i,β_i), then an application of a theorem of Bickel and Rosenblatt [3, p. 1073] and Theorem 3.3 yield the following result for the quadratic functional

$$U_n=(U_{1n},U_{2n},\ldots,U_{kn})$$

where

$$U_{in}=nb_n\int_{\alpha_i}^{\beta_{in}}[1-F(x)]^2[h_{in}^{(1)}(x)-h_i(x)]^2m_i(x)\,\mathrm{d}x$$

and m_i is a bounded piecewise smooth integrable function.

Corollary 4.2. *Let* $U=(U_1,\ldots,U_k)$ *with*

$$U_i=\int_{\alpha_i}^{\beta_i}\tilde{f}_i(x)m_i(x)\,\mathrm{d}x\int w_i^2(z)\,\mathrm{d}z.$$

Then, under the conditions (A1), (A2) *and* (A3), *with* $\tilde{f}_i$ *bounded away from zero on* (α_i,β_i), $b_n=n^{-\delta}$ $(\frac{2}{5}<\delta<\frac{1}{2})$ *and* $(1-F(\beta_{in}))^{-1}=\mathrm{O}(n^{-1/2-\lambda})$ *for some* $\lambda>0$, *we have that*

$$b_n^{-1/2}(U_n-U)$$

is asymptotically multivariate normal with independent components each having zero mean and variance

$$2\int\left\{\int w_i(x+y)w_i(x)\,\mathrm{d}x\right\}^2\mathrm{d}y\int m_i^2(x)\tilde{f}_i^2(x)\,\mathrm{d}x.$$

As stated in the remarks following Theorem 3.1, one can obtain certain strong laws for $\bar{T}_n$ via Theorem 3.1.

Révész [22, Theorem 4] has shown that

$$\lim_{n\to\infty} (2\lambda(w_i)\log b_n^{-1})^{-1/2} \sup_{\varepsilon<x<1-\varepsilon} |G_i(x,n)| \overset{\text{a.s.}}{=} 1$$

for any $\varepsilon > 0$, if the following conditions are satisfied:

(A4) (i) $\bar{f}_i$ is vanishing outside [0, 1],
(ii) $\bar{f}_i$ is twice differentiable in (0, 1) and $|\bar{f}_i''| \leq C$,
(iii) $\bar{f}_i(x) \geq \alpha > 0$ for all $x \in (0, 1)$,
(iv) w_i is a bounded symmetric density function which vanishes outside an interval $-\infty \leq A_i < B_i \leq \infty$,
(v) w is twice differentiable on (A_i, B_i) and $|w_i''| \leq C$,
(vi) $\lim_{x\to\infty} x^4 w_i(x) = 0$,
(vii) $\{b_n\}$ is a sequence of positive numbers satisfying $b_n \downarrow 0$, $nb_n \to \infty$,
(viii) $(\log b_n)^4/(nb_n \log b_n^{-1}) \to 0$, $nb_n^3/(\log b_n^{-1}) \to 0$.

Conditions (A4) are sufficient to prove the following version of (4.2),

$$\sup_{t\in R(\varepsilon)} \|\bar{T}_n(t) - G(t,n)\| \overset{\text{a.s.}}{=} o((\log b_n^{-1})^{1/2})$$

where $R(\varepsilon) = \{t \in \mathbb{R}^k : \varepsilon < t_i < 1-\varepsilon\}$, and hence

$$\lim_{n\to\infty} \bar{S}_n \overset{\text{a.s.}}{=} (1, 1, \ldots, 1) \tag{4.3}$$

where $\bar{S}_n = (\bar{S}_{1n}, \bar{S}_{2n}, \ldots, \bar{S}_{kn})$ and

$$\bar{S}_{in} = (2\lambda(w_i)\log b_n^{-1})^{-1/2} \sup_{\varepsilon<x<1-\varepsilon} |\bar{T}_{in}(x)|.$$

Since, in the present model, the G_i are independent Gaussian processes, the limit of $\bar{S}_n$ as $n \to \infty$ can be taken componentwise.

5. Asymptotic behavior of $H_n^{(2)}$

Consider the process

$$T_{in}^{(2)}(x) = (1-F(x))(nb_n/\bar{f}_i(x))^{1/2}\{h_{in}^{(2)}(x) - h_i(x)\} \tag{5.1}$$

for each i, $i = 1, 2, \ldots, k$, where $h_{in}^{(2)}$ is defined by (1.7) and h_i by (1.1). We will show under certain regularity conditions that $T_{in}^{(2)}$ and T_{in}, defined by (3.3), are asymptotically equivalent and hence prove a version of Theorem 4.1 for the vector process

$$T_n^{(2)}(t) = (T_{1n}^{(2)}(t_1), T_{2n}^{(2)}(t_2), \ldots, T_{kn}^{(2)}(t_k)) \tag{5.2}$$

where $t = (t_1, t_2, \ldots, t_k) \in \mathbb{R}^k$ and a version of Corollary 4.2.

We will assume, in addition to (A1), (A2) and (A3),

(A5) (i) the functions w_i vanishes outside the interval $[-A_i, A_i]$ where A_i is a finite positive number,
(ii) h_i, h_i' and h_i'' are all uniformly bounded on (α_i, β_i).

Theorem 5.1. *Under the conditions* (A1), (A2), (A3) *and* (A5) *we have*

$$\sup_{t\in R(n)} \|T_n^{(2)}(t)-T_n^*(t)\| \overset{\text{a.s.}}{=} O(r_5(n))$$

where $T_n^*(t)=(T_{1n}^*(t_1),\ldots,T_{kn}^*(t_k))$, (*cf.* (3.5)), $R(n)=\prod_{i=1}^k(\alpha_{in},\beta_{in})$ *and*

$$\begin{aligned} r_5(n)=\max\{&a_n(b_n\log\log n)^{1/2}[1-F(\beta_{in})]^{-1},\\ &v(n)b_n^{-1/2}n^{-1/3}[1-F(\beta_{in})]^{-2}(\log n)^{5/2},\\ &v(n)b_n^{-1/2}n^{-1/2}[1-F(\beta_{in})]^{-3}(\log n)^2\}\end{aligned}$$

with

$$v(n)=\max_{1\le i\le k} v_i(n) \quad \textit{and} \quad v_i(n)=\sup_{\alpha_{in}<x<\beta_{in}}(1-F(x))(\tilde{f}_i(x))^{-1/2},$$

and hence

$$\sup_{t\in R(n)} \|T_n^{(2)}(t)-G(t,n)\| \overset{\text{a.s.}}{=} O(\max\{r_1(n), r_2(n), r_5(n)\})$$

where $r_1(n)$ *and* $r_2(n)$ *are defined in* (3.4) *and* (3.6), *respectively.*

Remark. The rate of convergence in Theorem 5.1 is much poorer than that in Theorem 4.1. To get convergence to zero, we must choose the intervals $(\alpha_{in}, \beta_{in})$ so that

(i) $a_n[1-F(\beta_{in})]^{-1}=O(n^{\delta/2-\varepsilon})$ for some $\varepsilon>0$,
(ii) $v_i(n)[1-F(\beta_{in})]^{-2}=O(n^{1/3-\delta/2-\varepsilon'})$ for some $\varepsilon'>0$, and
(iii) $v_i(n)[1-F(\beta_{in})]^{-3}=O(n^{1/2-\delta/2-\varepsilon''})$ for some $\varepsilon''>0$.

Note that $\frac{1}{2}\delta<\frac{1}{4}$, $\frac{1}{3}-\frac{1}{2}\delta<\frac{1}{12}$ and $\frac{1}{2}-\frac{1}{2}\delta<\frac{1}{4}$.

Proof of Theorem 5.1. For each i, $1\le i\le k$, let

$$\begin{aligned} P_{in}(x)=&(1-F(x))(nb_n/\tilde{f}_i(x))^{1/2}\\ &\times\int b_n^{-1}w_i((x-u)/b_n)(1-F_n(u))^{-1}\,\mathrm{d}(\tilde{F}_{in}(u)-\tilde{F}_i(u))\end{aligned}$$

and

$$\begin{aligned} Q_{in}(x)=(1-F(x))(nb_n/\tilde{f}_i(x))^{1/2}\int b_n^{-1}w_i((x-u)/b_n)\\ \times(1-F(u))^{-1}\,d(\tilde{F}_{in}(u)-\tilde{F}_i(u)).\end{aligned}$$

The theorem follows from the following three lemmas where the conditions of the theorem are assumed.

Lemma 5.2.

$$\sup_{\alpha_{in}<x<\beta_{in}} |T_{in}^{(2)}(x) - P_{in}(x)| \overset{\text{a.s.}}{=} O(r_6(n)) \tag{5.3}$$

where

$$r_6(n) = \max\{v_i(n)n^{1/2}b_n^{5/2}, (b_n \log\log n)^{1/2} v_i(n)(1-F(\beta_{in}))^{-1}\}.$$

Proof. The left-hand side of (5.3) is bounded by

$$\sup_{\alpha_{in}<x<\beta_{in}} (1-F(x))(nb_n/\tilde{f}_i(x))^{1/2} \left| \int b_n^{-1} w_i((x-u)/b_n)(F_n(u)-F(u)) \times (1-F_n(u))^{-1} h_i(u)\, du \right| \tag{5.4}$$

$$+ \sup_{\alpha_{in}<X<\beta_{in}} (1-F(x))(nb_n/\tilde{f}_i(x))^{1/2} \left| \int b_n^{-1} w_i((x-u)/b_n) h_i(u)\, du - h_i(u) \right|.$$

As in Lemma 3.2 the second term of (5.4) converges to zero almost surely at the rate $O(v_i(n)n^{1/2}b_n^{5/2})$, while the first term is almost surely bounded by

$$(nb_n)^{1/2} v_i(n) \sup_{\alpha_{in}<x<\beta_{in}} \int_{-A_i}^{A_i} w_i(s)|F_n(x-b_ns) - F(x-b_ns)| \times (1-F_n(x-b_ns))^{-1} h_i(x-b_ns)\, ds$$

$$\overset{\text{a.s.}}{=} O((b_n \log\log n)^{1/2} v_i(n)[1-F(\beta_{in}+b_nA_i)]^{-1})$$

$$= O((b_n \log\log n)^{1/2} v_i(n)[1-F(\beta_{in})]^{-1})$$

by the boundedness of w_i and h_i, the law of the iterated logarithm applied to $(F_n - F)$, and since

$$[1-F_n(\beta_{in}+b_nA_i)]^{-1} \overset{\text{a.s.}}{=} O([1-F(\beta_{in}+b_nA_i)]^{-1})$$

and

$$[1-F(\beta_{in}+b_nA_i)]^{-1} = O([1-F(\beta_{in})]^{-1}).$$

The last statement follows from the boundedness of $f=F'$ and $b_n[1-F(\beta_{in})]^{-1} \to 0$ as $n \to \infty$.

Lemma 5.3.

$$\sup_{\alpha_{in}<x<\beta_{in}} |Q_{in}(x) - T_{in}^{*}(x)| \overset{\text{a.s.}}{=} O(r_7(n)) \tag{5.5}$$

where T_{in}^{*} *is defined by* (3.5) *and* $r_7(n) = a_n(b_n \log\log n)^{1/2}[1-F(\beta_{in})]^{-1}$.

Proof. The left-hand side of (5.5) is bounded by

$$\sup_{\alpha_{in}<x<\beta_{in}} (nb_n/\tilde{f}_i(x))^{1/2} \sup_{\alpha_{in}<x<\beta_{in}} \left| \int b_n^{-1} w_i(s) \right.$$
$$\left. \times [(1-F(x))(1-F(x-b_n s))^{-1}-1]\, \mathrm{d}(\tilde{F}_{in}(x-b_n s)-\tilde{F}_i(x-b_n s)) \right|$$
$$\leq (nb_n)^{1/2} a_n \sup_{\alpha_{in}<x<\beta_{in}} \left\{ \left| b_n^{-1} w_i(A_i) \left[\frac{1-F(x)}{1-F(x-b_n A_i)} - 1 \right] \right.\right.$$
$$\left. \times (\tilde{F}_{in}(x-b_n A_i) - \tilde{F}_i(x-b_n A_i)) \right|$$
$$+ \left| b_n^{-1} w_i(-A_i) \left[\frac{1-F(x)}{1-F(x+b_n A_i)} - 1 \right] \right.$$
$$\left. \times (\tilde{F}_{in}(x+b_n A_i) - \tilde{F}_i(x+b_n A_i)) \right|$$
$$+ \left| \int_{-A_i}^{A_i} (\tilde{F}_{in}(x-b_n s) - \tilde{F}_i(x-b_n s)) \right.$$
$$\times \left(w_i'(s) \left[\frac{1-F(x)}{1-F(x-b_n s)} - 1 \right] \right.$$
$$\left.\left.\left. + w_i(s) \left[\frac{1-F(x)}{1-F(x-b_n s)} h_i(x-b_n s) b_n \right] \right) \mathrm{d}s \right| \right\}. \tag{5.6}$$

Now, for $-A_i < s < A_i$,

$$\sup_{x<\beta_{in}} |(1-F(x))[1-F(x-b_n s)]^{-1} - 1| =$$
$$= \sup_{x<\beta_{in}} |F(x-b_n s) - F(x)| [1-F(x-b_n s)]^{-1}$$
$$\leq \sup_{x<\beta_{in}} b_n A_i f(x_n)[1-F(x-b_n s)]^{-1}$$
$$= \mathrm{O}(b_n [1-F(\beta_{in})]^{-1})$$

where $|x_n - x| \leq b_n A_i$. Since $f(x) = \sum_{i=1}^{k} \tilde{f}_i(x)$ by (A1)(ii), f is a bounded density. Consequently, the first two terms of the right-hand side of (5.6) converge almost surely to zero at the rate $\mathrm{O}(a_n b_n (\log\log n)^{1/2} [1-F(\beta_{in})]^{-1})$. Similarly, since $\int |w_i'(s)|\, \mathrm{d}s < \infty$, the first part of the integral in the last term of (5.6) converges almost surely to zero at the rate $\mathrm{O}(a_n b_n^{3/2} (\log\log n)^{1/2} [1-F(\beta_{in})]^{-1})$. For the final term,

$$a_n (nb_n)^{1/2} \sup_{\alpha_{in}<x<\beta_{in}}$$
$$\left| \int (\tilde{F}_{in}(x-b_n s) - \tilde{F}_i(x-b_n s)) w_i [1-F(x)][1-F(x-b_n s)]^{-1} \right.$$
$$\left. \times h_i(x-b_n s) b_n\, \mathrm{d}s \right|, \tag{5.7}$$

since $\sup_{x<\beta_{in}}(1-F(x))[1-F(x-b_ns)]^{-1}$ is uniformly bounded when $b_n(1-F(\beta_{in}))^{-1}\to 0$, we have that (5.7) converges almost surely to zero at the rate $O(a_n b_n^{5/2}(\log\log n)^{1/2})$, and hence the lemma.

Lemma 5.4.

$$\sup_{\alpha_{in}<x<\beta_{in}} |P_{in}(x)-Q_{in}(x)| \overset{\text{a.s.}}{=} O(r_8(n))$$

where

$$r_8(n)=v_i(n)b_n^{-1/2}\max\{n^{-1/3}[1-F(\beta_{in})]^{-2}(\log n)^{5/2},\; n^{-1/2}[1-F(\beta_{in})]^{-3}(\log n)^2\}.$$

Proof. The proof is similar to that of [5, Theorem 4.2], in particular, the handling of $R'_{4n}(t)$.

6. Concluding remarks

The results of Theorem 4.1 concerning the statistic $\bar{M}_n$ are useful for constructing approximate confidence bands for the hazard rate vector H of (1.5). One can substitute F_n for F and $\tilde{f}_{in}^{-1/2}$ for $\tilde{f}_i^{-1/2}$ where they appear as factors in the definition of $\bar{T}_{in}$, cf. (4.1). The results for $\bar{M}_n$ and U_n enable one to test the hypothesis that $H(t)=H(t;\theta)$, where θ is a vector of unknown parameters by estimating θ by a suitable sequence $\hat{\theta}_n$ of estimators and substituting the estimator $\hat{\theta}_n$ for the unknown value θ_0 of θ where it appears in $\bar{M}_n$ and U_n. The resulting statistics would have the same asymptotic distribution as in the case when all parameters are specified, provided, for example, that $\|\hat{\theta}_n-\theta_0\|=O(n^{-1/2})$ as is usually the case for maximum likelihood estimators. Local power calculations can also be made. We refer to Bickel and Rosenblatt [3] for further details on these statistical considerations.

Consider the sub-system $\{Y_i^j\}_{i\in J}$ where $\emptyset\neq J\subset\{1,2,\ldots,k\}$ as described in Section 2. If one assumes condition (2.4), then F_J of (2.2) can be consistently estimated by the Kaplan–Meier [13] estimator $\hat{F}_{Jn}$ defined by

$$1-\hat{F}_{Jn}(x)=\begin{cases}\prod_{j:X_j<x}[(n-R_j)/(n-R_j+1)]^{\sum_{i\in J}\delta_j^i} & \text{if } x<X_{(n)},\\ 0 & \text{otherwise}\end{cases}$$

where R_j is the rank of X_j and $X_{(n)}=\max\{X_1,X_2,\ldots,X_n\}$. Thus, one can estimate h_J of (2.3) by

$$(1-\hat{F}_{Jn}(x))^{-1}\int b_n^{-1}w_J((x-u)/b_n)\,\mathrm{d}\hat{F}_{Jn}(u) \tag{6.1}$$

or by

$$\int b_n^{-1}w_J((x-u)/b_n)(1-\hat{F}_{Jn}(u))^{-1}\,\mathrm{d}\hat{F}_{Jn}(u)$$

where w_J is a kernel satisfying (A2). Földes, Rejtö and Winter [11] have obtained strong consistency results for (6.1). The corresponding approach in this paper is to estimate (2.3) by

$$\sum_{i \in J} h_{in}^{(1)}(x) \quad \text{or} \quad \sum_{i \in J} h_{in}^{(2)}(x) \tag{6.2}$$

where $h_{in}^{(1)}$ and $h_{in}^{(2)}$ are defined by (1.5) and (1.7), respectively. The latter estimators are simpler than (6.1), since $d\tilde{F}_{in}$ of (1.3) and (1.7) gives constant mass $1/n$ at each order statistic $X_{(m)}$ for which $\delta^i_{(m)} = 1$, while $d\hat{F}_{Jn}$ of (6.1) gives random mass $(j_{m-1} - j_m)$ at each $X_{(m)}$ for which $\sum_{i \in J} \delta^i_{(m)} = 1$, where

$$j_m = \prod_{j=1}^{m} [(n-j)/(n-j+1)]^{\sum_{i \in J} \delta^i_{(j)}}$$

and $\delta^i_{(m)}$ the corresponding indicator function for $X_{(m)}$. The quantity depends on which of the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(m-1)}$ are due to a failure of sub-system J and which are due to a failure of J^c. As was shown in this paper the components of (1.6) and (1.8) are asymptotically independent so that tests for one or more components of the hazard rate vector H of (1.2) would be asymptotically independent from tests for the other components. A comparison between Kaplan–Meier-type density or hazard rate estimation (6.1) and that of the present exposition will be carried out in a subsequent paper.

The Kaplan–Meier estimator is also used in the estimation of the cumulative hazard rate (log survivor function). A number of studies have been done on this problem and we refer to the treatments of Efron [9], Breslow and Crowley [4], Aalen [1], Burke, Csörgö and Horváth [5] and to Csörgö and Horváth [8].

As to the maximum deviation of empirical density estimator, we refer to Konakov [16] where complete asymptotic expansions are obtained.

Acknowledgement

I wish to thank the referees for their helpful comments on the original manuscript and E. Arjas for a helpful conversation.

Note added in proof

After this manuscript was sent to the typesetter, I became aware of [29], in which the limiting distributions of $\bar{M}_{1n}$ and U_{1n} are obtained in the case $k = 2$ assuming $\{Y_j^1\}$ and $\{Y_j^2\}$ are independent. In the present paper, written independently of [29], the limiting *joint* distributions of the vectors M_n and U_n are obtained and the above independence condition is not assumed.

References

[1] O. Aalen, Nonparametric inference in connection with multiple decrement models, Scand. J. Statist. 3 (1976) 15–27.

[2] E. Arjas, A note on multiple causes of failure and the Kaplan–Meier and Nelson plots, 1981.

[3] P.J. Bickel and M. Rosenblatt, On some global measures of the deviation of density function estimates, Ann. Statist. 1 (1973) 1071–1095.

[4] N. Breslow and J. Crowley, A large sample study of the life table and product limit estimates under random censorship, Ann. Statist. 2 (1974) 437–453.

[5] M.D. Burke, S. Csörgö and L. Horváth, Strong approximations of some biometric estimates under random censorship, Z. Wahrsch. Verw. Gebiete 56 (1981) 87–112.

[6] D.R. Cox, Renewal Theory (Methuen, London, 1962).

[7] D.R. Cox and P.A.W. Lewis, The Statistical Analysis of Series of Events (Methuen, London, 1966).

[8] S. Csörgö and L. Horváth, On cumulative hazard processes under random censorship from the right, Preprint, 1980.

[9] B. Efron, The two-sample problem with censored data, Proc. 5th Berkeley Symp. Probab. Statist. 4 (1967) 831–853.

[10] V.A. Epanechnikov, Nonparametric estimates of a multivariate probability density, Theory Probab. Appl. 14 (1969) 153–158.

[11] A. Földes, L. Rejtö and B.B. Winter, Strong consistency properties of nonparametric estimators for randomly censored data Part II: Estimation of density and failure rate, Preprint, 1980.

[12] V.N. Gaisky, On the time distribution of large, deep earthquakes from the Pamir–Hindu–Kush, Physics of Solid Earth 6 (1966) 343–349 (English translation).

[13] E.L. Kaplan and P. Meier, Nonparametric estimation from incomplete observations, J. Amer. Statist. Assoc. 53 (1958) 457–481.

[14] A.W. Kimball, Estimation of mortality intensities in animal experiments, Biometrics 16 (1960) 505–521.

[15] J. Komlós, P. Major and G. Tusnády, An approximation of partial sums of independent r.v.'s and the sample d.f. Part I, Z. Wahrsch. Verw. Gebiete 32 (1975) 111–132.

[16] V.D. Konakov, Complete asymptotic expansions for the maximum deviation of the empirical density function, Theory Probab. Appl. 23 (1978) 475–489.

[17] N. Langberg, F. Proschan and A.J. Quinzi, Converting dependent models into independent ones, preserving essential features, Ann. Probab. 6 (1978) 174–181.

[18] N. Langberg, F. Proschan and A.J. Quinzi, Estimating dependent life lengths, with applications to the theory of competing risks, Ann. Statist. 9 (1981) 157–167.

[19] A.V. Peterson, Nonparametric estimation in the competing risks problem, Ph.D. Thesis, Stanford University, 1975.

[20] R.L. Prentice, J.D. Kalbfleisch, A.V. Peterson, V.T. Farewell and N.E. Breslow, The analysis of failure times in the presence of competing risks, Biometrics 34 (1978) 541–554.

[21] P. Révész, On empirical density function, Periodica Math. Hungar. 2 (1972) 85–110.

[22] P. Révész, A strong law of the empirical density function, Periodica Math. Hungar. 9 (1978) 317–324.

[23] J. Rice and M. Rosenblatt, Estimation of the log survivor function and hazard function, Sankhyā A 38 (1976) 60–78.

[24] M. Rosenblatt, Curve estimates, Ann. Math. Statist. 42 (1971) 1815–1842.

[25] A. Tsiatis, A nonidentifiability aspect of the problem of competing risks, Proc. Nat. Acad. Sci. U.S.A. 72 (1975) 20–22.

[26] A. Udias and J. Rice, Statistical analysis of microearthquake activity near San Andreas Geophysical Observatory, Hollister, California, Bull. Seismological Soc. Amer. 65 (1975) 809–828.

[27] G.S. Watson and M.R. Leadbetter, Hazard analysis Part I, Biometrika 51 (1963) 174–184.

[28] G.S. Watson and M.R. Leadbetter, Hazard analysis Part II, Sankhyā A 26 (1964) 110–116.

[29] J.R. Blum and V. Susarla, Maximal deviation theory of density and failure rate function estimates based on censored data, in: P.R. Krishnaiah, ed., Multivariate Analysis Vol. V (North-Holland, Amsterdam, 1980) pp. 213–222.